

A Social Psychologist Encounters the Real World 189

Invited Address

Annual Meeting of the American Psychological Association
Washington, D.C. August, 1986

Robert L. Helmreich
The University of Texas at Austin

The study of group behavior and performance, while at the core of social psychology, has proved to be one of the more recalcitrant areas of investigation. A host of methodological problems and threats to external validity beset the researcher at every turn. These range from attempting to capture the processes involved in multi-person interactions, to generalizing from groups of strangers briefly assembled performing relatively meaningless tasks without long-term consequences, to partitioning individual versus group variance in performance, and a litany of other vexations.

Looking at the state of research in the area, Ivan Steiner (Steiner, 1973) felt constrained to write an article in 1973 entitled "Whatever happened to the group in social psychology? What happened to the group was a massive defection to other, more tractable research problems. Indeed, despite early involvement in group research and continuing interest, I also must be counted among the defectors.

For a number of years the main thrust of my research, much of it in collaboration with Janet Spence, has been on isolating components of personality and achievement motivation that determine performance, particularly in demanding work. We have

achieved a fair amount of success in this endeavor, which now includes reformulations of achievement motivation as a three component construct and of the Type A personality constellation. In these investigations we have looked at performance in groups as diverse as school children, college students, businessmen and scientists (Helmreich & Spence, 1978; Spence & Helmreich, 1983). The logical but unaccomplished extension of this work was to the performance of groups.

It was in 1979, when the work on achievement had just begun to achieve coherence, that serendipity struck in the form of John K. Lauber. At that time John was an experimental psychologist working in aviation human factors at Ames Research Center. (He is now the first psychologist sitting as a member of the National Transportation Safety Board.) Having found some of my work potentially relevant to issues of crew performance, he invited me to participate in a NASA/Airline conference on crew coordination. In the process of preparing for the meeting, I discovered the incredible research potential of the flightdeck. Here was a structured environment with well defined goals and well defined roles where small groups had to perform in a highly coordinated manner, to evaluate large amounts of information, and to make decisions with life or death import.

Several other considerations increased the appeal of the flightdeck as a research setting. The first was that the flight situation, while open to many variations, has enough commonalities to make differences in performance observable and measurable. The second came from looking at the results of

research from NASA and the National Transportation Safety Board (Cooper, White, & Lauber, 1979). Analyses of the causes of jet transport accidents over a period of many years led to several clear and conclusions: one being that more than two-thirds of all accidents result from "pilot error" rather than any failure in the aircraft or catastrophic environmental condition. More significantly, again in a majority of accidents, the slippery term "pilot error" did not refer to mistakes in manipulation of controls or lack of technical competence but rather to failures in crew coordination and management of the resources available to the flightdeck.

These data point to a social psychological research setting where critical outcomes grow inexorably out of group processes. Of course, the overwhelming majority of aircraft flights do not end in crashes caused by poor group process. However, it is possible to recreate the conditions surrounding an accident or incident in the flight simulator. And that was the third consideration in choosing to study flight crews. The modern simulator has astonishing fidelity, recreating all aspects of flight from the view out the window, aircraft motion and instrument readings, to the pitter-patter of raindrops hitting the fuselage. Clay Foushee has described the research potential of the simulator in his well-titled American Psychologist article, "Dyads and triads at 35,000 feet" (Foushee, 1984). It is important to note that professional flight crews view the simulator as an integral part of their flying experience rather than a research tool. Crews are now trained in the simulator rather than the aircraft and their performance for purposes of

maintaining their licensure is evaluated regularly in "the box".

Finally, there were more practical and personal considerations, one of them being the intense interest and commitment of the aviation community to any research that might improve safety and operational efficiency. It is rare and gratifying to find consumers of research who are eager to employ the results of research. In the same vein, it is both challenging and rewarding to conduct research that might not only provide theoretical insights but also prove directly and immediately beneficial in the real world. Given a commitment to conducting research in this arena, it was possible (and naive) to construct a meta-plan for how research would be organized and would progress. In its pristine simplicity, the plan strategy involved the collection of the relevant personality, attitudinal, and demographic data on large samples of pilots and correlating these with hard performance data collected both in natural (line) settings and in the simulator. In addition to provide valuable insights on the relationships among personality, attitudinal, and performance factors, the envisioned data would also provide guidance for pilot selection and assist in the evaluation the effectiveness of new training programs designed to improve crew coordination and cockpit resource management.

It was at this point that the realities of the "real world" became most evident. Some problems were generic, including the financial pressures of deregulation on organizations and the loss of one participating airline to merger. Other problems related more to the nature of the research. While three major elements in

civil air transport, airline management, the Federal Aviation Administration, and pilots themselves, all enthusiastically endorse research in the interests of improved safety, there are powerful forces that hinder investigation. From the perspective of the pilots, there is apprehension that research on pilot performance might result in data which could be used either by management or by the FAA to jeopardize their licenses and livelihood. From the perspective of the FAA, there is an obligation to maintain the highest level of safety, which means that any data reflecting sub-standard performance or procedures should be used to eliminate potential risk - a deadly threat to data collection. Finally, there is a real threat to airlines themselves. If extensive data exist on pilot performance and a pilot from the lower end of the distribution were to be involved in a major accident, the potential liability to the organization is staggering. Current evaluation practices in aviation raise yet another hurdle. While operating a modern jet transport is undeniably a group endeavor, Federally mandated performance evaluations focus almost exclusively on individual performance and the ability to use stick and rudder in such evolutions as making steep turns and aborting take-offs at high speed. The required annual Proficiency Check is geared toward the evaluation of these technical aspects of flying. What is lacking is systematic evaluation of the group actions that research has shown to be the primary cause of accidents. What is also lacking is a research literature on crew performance in operational settings. Research on pilot performance has been restricted almost exclusively to performance in training and has

concentrated on individual technical proficiency. Both in research and practice, the tradition of the Red Baron endures. There are also substantial differences between air carriers in pilot characteristics, attitudes, and practices making organizational climate an important element in the research (Hackman, 1986; Helmreich, 1986). Richard Hackman, Clay Foushee and I have discussed these issues in more detail in a recent paper (Helmreich, Hackman, & Foushee, 1986).

The practical effect of this assortment of obstacles is that progress has been much slower than hoped for and the starting point much more basic than anticipated. However, we have now reached a position of trust and cooperation and with several organizations regarding the sanctity and confidentiality of data and have begun to assemble a massive, protected database with information from a number of major airlines and elements of the Air Force, Military Airlift Command. The data on hand, although less extensive than we would like nonetheless have produced some highly encouraging results. During the remainder of the discussion I will describe some of our findings and where we are heading in the future.

Given the tradition of individual evaluation and needing to work within the current system, our first step was to try to develop more comprehensive measures of individual performance that took into account both technical proficiency and the managerial, interpersonal aspects of flightdeck behavior. This effort was greatly facilitated by the research of Lauber, White, Cooper, Ruffell Smith, and their colleagues at NASA (Cooper,

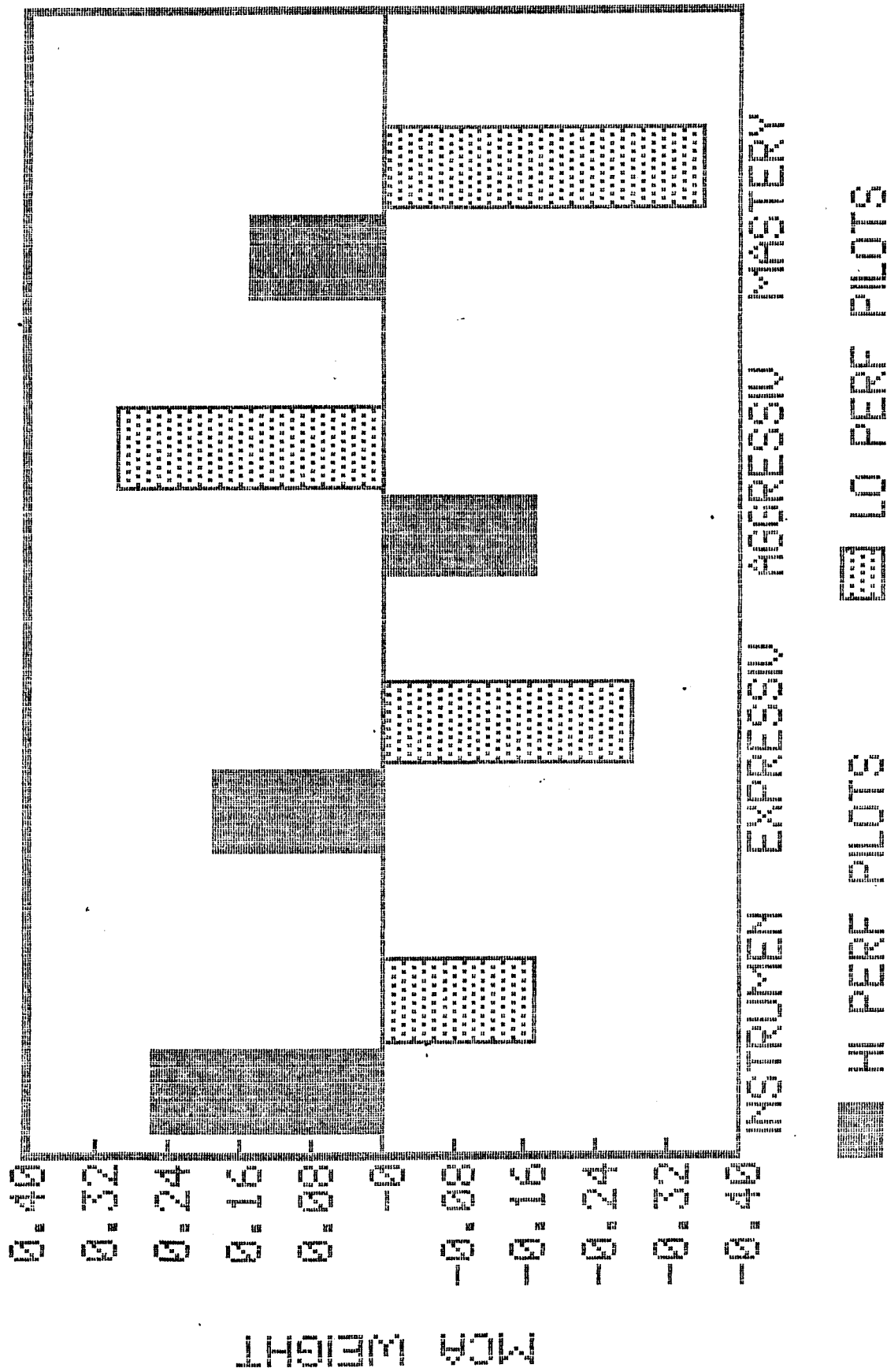
White, & Lauber, 1979 Ruffell Smith, 1979). This work, conducted as controlled simulator research, structured interviews with line pilots, and analyses of aircraft accidents, had isolated both behaviors and attitudes associated with effective cockpit management.

Working with Check Airmen (highly qualified airline pilots who are qualified by the FAA to conduct mandatory evaluations of line and simulator performance of pilots) at a cooperating airline, we have been able to obtain highly reliable ratings of both the technical and managerial aspects of flight operations. The results show, not surprisingly, that the two dimensions of performance are positively correlated but that there is substantial unshared variance. (The average correlation between technical proficiency and resource management is around 0.40.) We have also been able to administer our two primary personality measures, the Extended Personal Attributes Questionnaire (EPAQ: Spence & Helmreich, 1978; Spence, Helmreich, & Holahan, 1979), and the Work and Family Orientation Questionnaire (WFOQ: Helmreich & Spence, 1978) to a large number of pilots. The EPAQ measures positive and negative constellations of instrumental and expressive traits while the WFOQ assesses three facets of achievement motivation, Mastery Needs, Work Orientation, and interpersonal Competitiveness.

We now have data from three studies in two organizations relating personality to performance (Helmreich, 1982; 1986). The results show consistent, significant relationships between traits and line performance. Figure 1 shows some of the more potent factors that discriminate between airline pilots globally

ORIGINAL PAGE IS
OF POOR QUALITY

Fig. 1: PILOT PERSONALITY AND PERFORMANCE



classified as high or low performing by Check Airmen. Instrumentality, Expressivity, and high Mastery needs characterize the better pilots while high Aggressiveness is an attribute of the low rated group. More recently, Spence and I, with graduate student Robby Pred, have reformulated the Type A personality construct. In our new, factor analytically derived measure, there are two moderately correlated scales we have called Driveness and Impatience. The most interesting aspect of this model is that high Driveness appears to be associated with superior performance and to have no negative health implications, while high Impatience has both negative performance and negative health relationships. In his recent dissertation, Tom Chidester (1986) included the new Type A measure as part of a study of flightcrew reactions and performance in long and short-haul operations. He found Driveness to be positively correlated with resource management and uncorrelated with technical performance. Impatience, on the other hand, was negatively correlated with technical performance and uncorrelated with resource management.

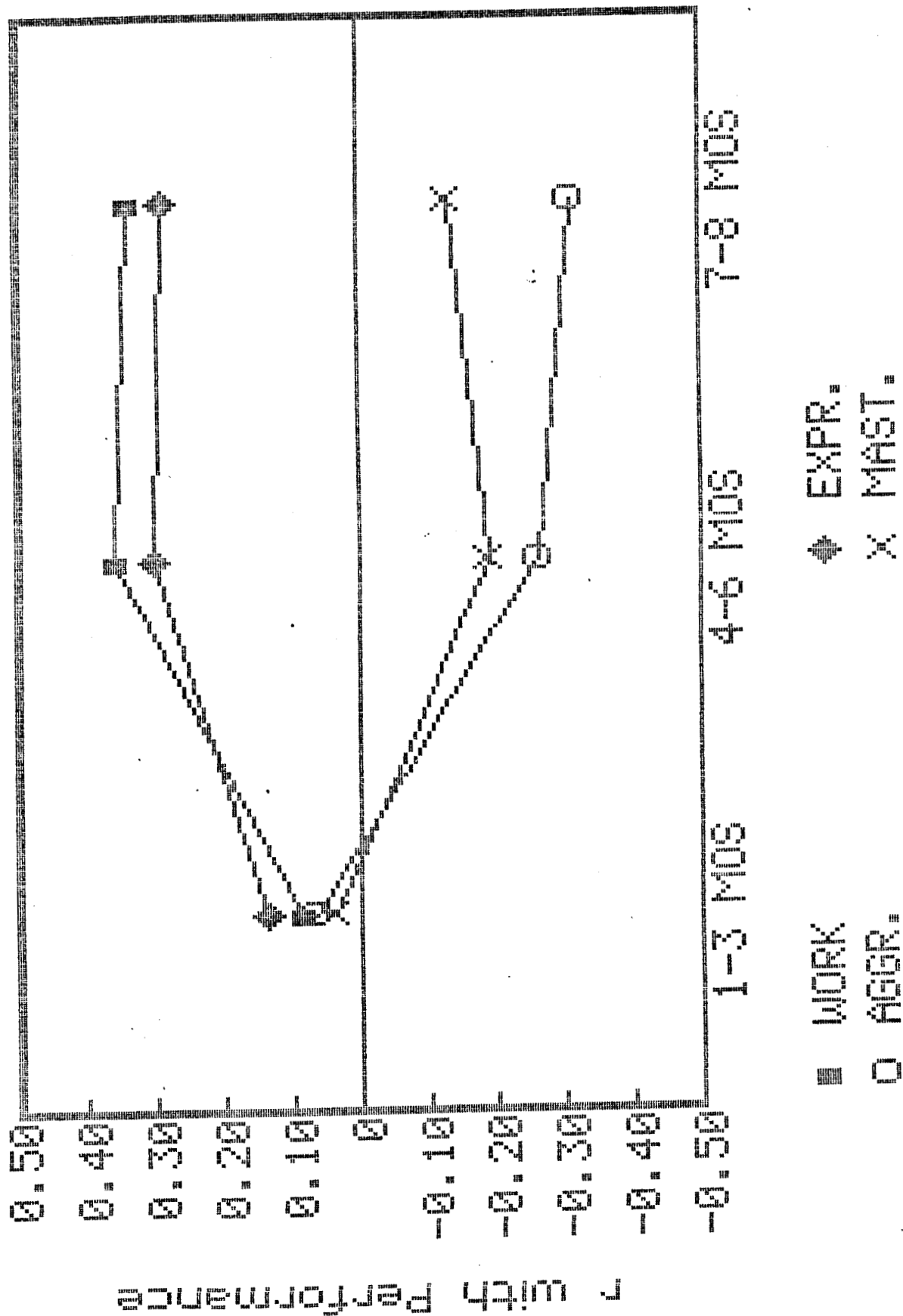
It is gratifying to find personality to be a robust predictor of real world performance, given its rather tenuous place in the world of social psychology. It is even more striking to find it operating in a situation where many other factors ranging from personal attitudes to group composition also influence performance.

The obtained personality-performance relationships are also at variance with the research literature on pilot selection. Reviewing this literature from World War II onward, one finds

little evidence for personality effects. Indeed, a review of research on Navy pilot selection conducted in the late 1970's concluded that there was no evidence for any personality factors being related to pilot attrition during training (Griffin & Moskos, 1977). Why should we have found consistent personality effects when others have had no success? We feel that the answer to this seeming paradox lies in what we have labelled the "honeymoon effect" of motivation on performance. In a recent study (Helmreich, Sawin, & Carsrud, 1986) we examined the correlations between our achievement scales and performance over time in a sample of clerical workers. Figure 2 shows the effect graphically. At the end of training there were no significant correlations. However, with the passage of time, the correlations increased in magnitude and stabilized. Our theoretical interpretation is that most individuals, when selected for a position, give their utmost effort during training and/or initial work period and that this high effort masks the influence of personality. It is not until the person has settled into the routine of the position and the "honeymoon has ended" that personality influences on behavior begin to emerge. This delayed impact may also explain why personality effects are frequently elusive in the traditional laboratory study.

As noted earlier, NASA research had isolated certain pilot attitudes that appeared to be associated with effective or ineffective cockpit management. As part of the exploration of pilot behavior, I decided to investigate these attitudes more systematically, thus returning to my first graduate school research topic. Using the available data as a resource, I

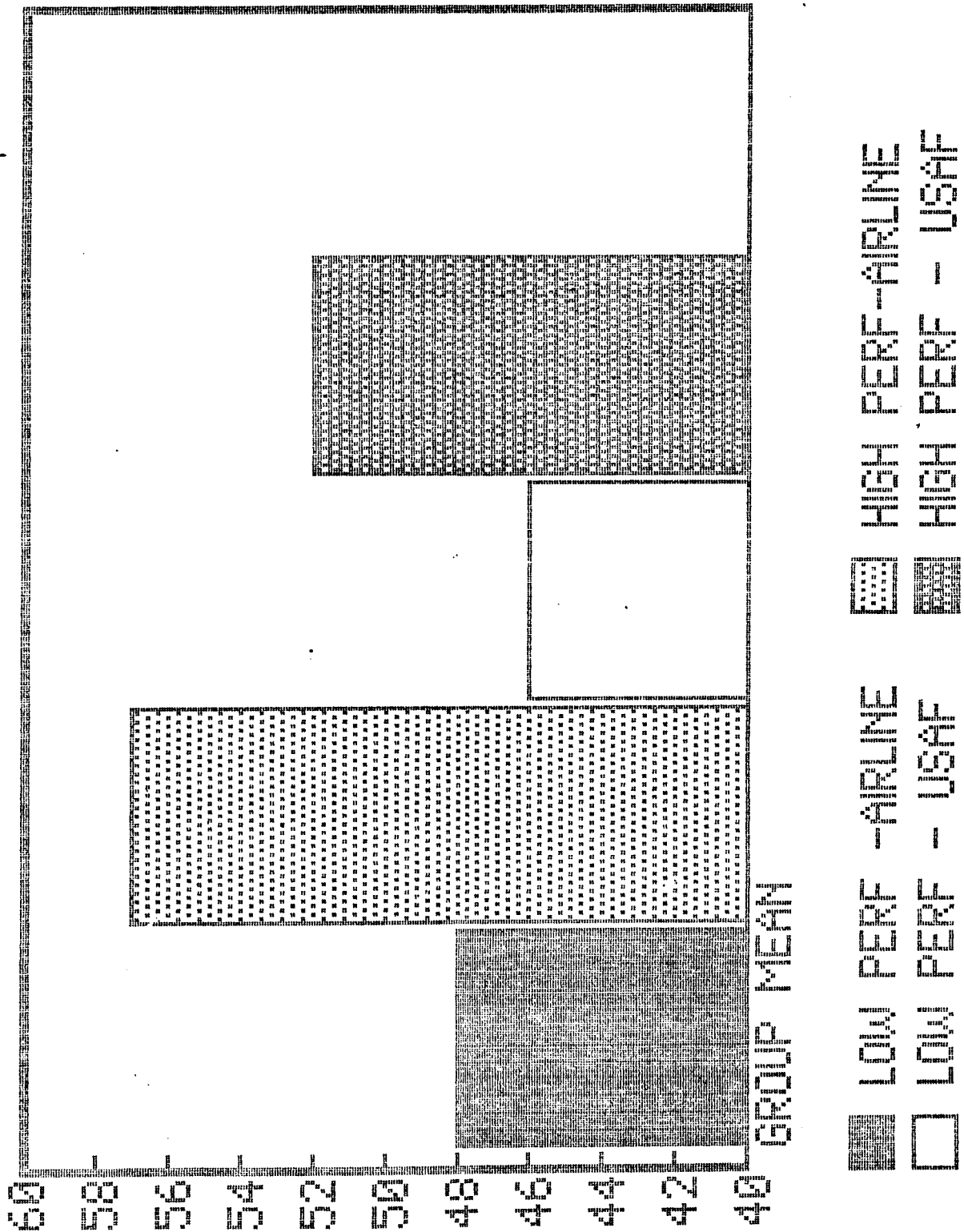
Fig. 2: PERSONALITY AND PERFORMANCE



developed a twenty-five item "Cockpit Management Attitudes Survey. (Helmreich, 1984). This instrument has now been administered to more than 5,000 pilots from four major airlines, the US Air Force, and the general aviation community. The questions cover a range of topics from attitudes about personal reactions and decision making under stressful conditions, to crew roles and responsibilities, interpersonal communications, and training. The results have proved quite striking. A recent study validated the attitude-performance linkage. Using Check Airmen who were blind to pilot attitudes as evaluators, two groups of pilots were identified. Both groups were high in technical proficiency, but one group was classified as high on resource management and the other as low. Using discriminant function analysis, 94% of the pilots were correctly classified as to condition (Helmreich, Foushee, Benson, & Russini, 1986). Figure 3 shows group differences on the composite attitude index in the airline and in a replication just completed with Air Force transport pilots. There seems to be little doubt that in the real world attitudes and behaviors are clearly linked.

There are several other intriguing aspects of the attitude data. The first is that there are large and highly significant differences between organizations in attitude. This seems to verify Hackman's (1986) observation that organizational culture and climate are prime determinants of operational attitudes and performance. Even more striking are large and significant differences in attitude between crew position (i.e Captain, First Officer, Flight Engineer) within organizations. A clear

Fig 3: PILOT ATTITUDES AND PERFORMANCE



implication of this finding is that if crewmembers on the same aircraft do not agree as to how best to manage the flight operation, the level of effective teamwork is likely to be less than optimal. Investigating these attitude differences, one of our graduates students, Valerie Edwards did separate cluster analyses of the attitudes of Captains, First Officers, and Flight Engineers (Edwards, 1986). She found that the attitudes of Captains formed more logically consistent clusters, while those of the other crewmembers tended to be more fractionated.

A reassuring implication of these data is that since attitudes can be changed, training programs directed towards the modification of attitudes should be able to produce more positive attitudes, greater agreement among crewmembers, and more effective performance. Several airlines have instituted formal training in cockpit resource management and we are attempting to assess the short and long-term effects of such programs on both attitudes and performance. Although the data are just beginning to come in, preliminary indications are that attitudes are significantly changed by participation in this type of training. Performance effects and the stability of changes remain to be determined.

While the results to date are encouraging and provide strong inducements to continue this line of investigation, a number of critical questions remain unanswered. These include gaining more understanding of organizational effects, further definition of personality dimensions relevant to performance and to pilot selection, and exploring relationships between attitudes and personality, as well as defining the limits of training in crew

coordination. Additionally, as modern aircraft become increasingly automated and are operated with two human and an electronic crewmember, it will become imperative to study the social psychological impact of automation on crew attitudes and performance and to discover whether the optimum pilot personality constellation changes with the changing environment of the aptly labelled "glass cockpit".

What has been missing from the research and data thus far are references to the group, the point on which I opened the discussion. This omission does not reflect abandonment of the initial commitment to study flight crews as small groups. Rather, it reflects the need to understand the setting and the individual players before trying to disentangle the interpersonal interactions that make up the conduct of flight operations. At the same time that we have been investigating the aspects of pilot behavior just described, we have been trying to develop instruments and methodologies to capture the group processes and communications involved in normal and abnormal flight situations. Paralleling this effort, Clay Foushee and his colleagues at Ames have made great progress in the analysis of pilot behavior using real crews flying a standardized flight scenario involving weather and equipment problems in the simulator (Foushee, Lauber, Baetge, & Acomb, 1986; Foushee & Helmreich, in press). Lauber, 1986). Richard Hackman and his colleagues have been systematically observing crew behavior in regular flight operations. In our own laboratory, Rick Siem has conducted an ingenious dissertation using volunteer crews flying a micro-

computer based flight simulator. Siem is in the process of analyzing crew interactions coded from videotapes of the complete flights of each crew.

We feel that we are now ready to tackle the group rather than the individual on the flightdeck. What is developing under NASA sponsorship is a synergistic relationship among a group of investigators including Foushee and his colleagues at NASA, Hackman's team from Harvard, and our group at Texas. The research of each has both common and unique dimensions. Our group, for example, will be collecting systematic observational data in both simulators and on line flights as well as attitude and personality data.

At the same time, all of us will collaborate on a large study using the highly sophisticated research simulator at NASA Ames which has elaborate data acquisition facilities as well as state of the art fidelity and the capacity to simulate the air control environment. The planned study will involve the collection of detailed data on technical performance and crew interaction and will also investigate the effects of crew composition along selected personality dimensions. The database resulting from this collaboration should allow us to explore a number of hypotheses about effective and ineffective behavior and how individual attributes influence group process.

In conclusion, I must admit that research on flightcrew behavior remains up in the air. But if we can keep it there, I am becoming increasingly optimistic about our chances of gaining insights of both theoretical and practical import.

References

- Chidester, T.R. (1986) Mood, sleep, and fatigue effects in flight operations. Unpublished doctoral dissertation. The University of Texas at Austin.
- Cooper, G.E., White, M.D. & Lauber, J.K. (Eds.), (1979) Resource Management on the flight deck. NASA Report No. CP-2120.
- Edwards, V.J. (1986) Flightdeck management attitudes: A cluster analysis. In G.E. Lee (Ed.) Proceedings of the Tenth Symposium: Psychology in the Department of Defence. Colorado Spgs: U.S. Air Force Academy.
- Foushee, H.C. (1984) Dyads and triads at 35,000 feet: Factors affecting group process and aircrew performance. American Psychologist, 39, 886-893.
- Foushee, H.C. & Helmreich, R.L. (in press) Group interactions and flightcrew performance. In E. Weiner & J. Nagel (Eds.), Human factors in modern aviation. New York: Academic Press.
- Foushee, H.C. Lauber, J.K., Baetge, N.M., & Acomb, D.B. (1986) Crew factors in flight operations III: The operational significance of exposure to short haul air transport operations. Moffett Field, CA: NASA-Ames Research Center.
- Griffin, G.R. & Mosko, J.D. (1977) Naval aviation attrition 1950-1976: Implications for the development of future research and evaluation. Pensacola, Florida: Naval Aerospace Medical Research Laboratory.
- Hackman, J.R. (1986)
- Helmreich, R.L. (1982) Pilot selection and training. Paper presented at the American Psychological Association, Annual Meeting, Washington, D.C.
- Helmreich, R.L. (1984) Cockpit management attitudes. Human Factors, 26, 583-589.
- Helmreich, R.L. (1986) Pilot selection and performance evaluation: A new look at an old problem. In G.E. Lee (Ed.) Proceedings of the Tenth Symposium: Psychology in the Department of Defence. Colorado Spgs: U.S. Air Force Academy.
- Helmreich, R.L., Hackman, J.R., & Foushee, H.C. (1986) Evaluating flightcrew performance: Policy, pressures, pitfalls, and promise. NASA/University of Texas Technical Report 86-1.
- Helmreich, R.L., Sawin, L.L. & Carsrud, A.L. (1986) The honeymoon effect in job performance: Delayed predictive power of achievement motivation. Journal of Applied Psychology.

Helmreich, R.L. & Spence, J.T. (1978) The Work and Family Orientation Questionnaire: An objective instrument to assess components of achievement motivation and attitudes toward family and career. JSAS Catalog of Selected Documents in Psychology, 8, 5, MS 1677.

Ruffell Smith, H.P. (1979) A simulator study of the interaction of pilot workload with errors, vigilance, and decisions. Moffett Field, CA. NASA-TM 78482.

Spence, J.T. & Helmreich, R.L. (1978) Masculinity and femininity: Their psychological dimensions, antecedents, and correlates. Austin: University of Texas Press.

Spence, J.T., Helmreich, R.L., & Holahan, C.J. (1979) Negative and positive components of psychological masculinity and femininity and their relationships to self-reports neurotic and acting out behaviors. Journal of Personality and Social Psychology, 37, 1673-1682.

Steiner, I.D. (1974) Whatever happened to the group in social psychology? Journal of Experimental Social psychology, 10, 94-108.

The author's research reported here and preparation of the report were supported by NASA Cooperative Agreement NAG2-137, Robert L. Helmreich, Principal Investigator. Special thanks are due to John Wilhelm for his comments on suggestions on the paper and collaboration at all stages of the research.